DOCUMENT RESUME

ED 222 584 TM 820 799

AUTHOR Adwere-Boamah, Joseph
TITLE Getting More Information from School District Surveys with Goodman's "Modified Regression Approach."
PUB DATE [80]
NOTE 18p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Attitude Measures; Elementary Secondary Education; Evaluation Methods; *Mathematical Models; *Maximum Likelihood Statistics; Predictive Measurement; *Predictor Variables; *Regression (Statistics); Statistical Analysis
IDENTIFIERS *Categorical Data; *Logit Analysis; Log Linear Models

ABSTRACT

The development of new statistical methods of log-linear models by Leo Goodman has led to major advances in the statistical analysis of categorical data. Goodman's logit analysis, (the simplest form of log-linear models) can be applied in evaluation studies to estimate the "main effects" and "interaction effects" of categorical explanatory variables on a dichotomous dependent variable. Logit analysis is analogous to multiple regression analysis of a continuous dependent variable. It is a "modified regression approach" in place of the linear model of multiple regression, and the estimation procedure is maximum likelihood. The logit model was used with hypothetical data to analyze and predict classroom teachers' attitudes toward proficiency tests based on dichotomous explanatory variables of ethnicity, teaching grade level, and sex of the teachers. The partial regression coefficients were calculated and used to find main effects and interaction effects. Additive and saturated models were used to find the independent logit model in which the criterion analyzed reflects the expected odds of attitude as a function of the explanatory variables. An analysis of the indices of the magnitude of contribution is provided. (CM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED222584

GETTING MORE INFORMATION FROM SCHOOL DISTRICT SURVEYS WITH GOODMAN'S "MODIFIED REGRESSION APPROACH"

JOSEPH ADWERE-BOAMAH

OAKLAND UNIFIED SCHOOL DISTRICT
Department of Research and Evaluation

TM 820 755

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Adwere-Boamah

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

GETTING MORE INFORMATION FROM SCHOOL DISTRICT SURVEYS WITH GOODMAN'S "MODIFIED REGRESSION APPROACH"

Program evaluators and school district people encounter qualitative or categorical data regularly in their activities. Many of the tasks concerned with program evaluation and planning involve systematic collection and analysis of qualitative data such as survey results of students', teachers', or parents' attitudes toward a particular program. Evaluation activities often seek to explain for example, variations in the attitudes or opinions of a target population of teachers in a school district toward a program. The traditional methods (e.g., percentages) for summarizing cross-tabulated data are inadequate, in many instances for answering evaluation questions. For example, when the focus of an evaluation study is on analysis and prediction of a dichotomous dependent variable (e.g., positive or negative attitude toward a program) as a function of other qualitative explanatory variables, the usual regression model is inappropriate for the analysis. The development of new statistical methods, the log-linear models by Leo Goodman has led to major advances in the statistical analysis of categorical data. They enable the analyst to identify simple and complex relationships between two or more cross-classified variables. For example, Goodman's logit analysis, (the simplest form of log-linear models) can be applied in evaluation studies to estimate the "main effects" and "interaction effects" of categorical explanatory variables on a dichtomous dependent variable.

Logit analysis is analogous to multiple regression analysis of continuous dependent variable. It is a "modified regression approach," in that the logit model is used in place of the usual linear model of multiple regression, and the estimation procedure is maximum likelihood instead of least squares. Despite the widespread collection of categorical data as a major part of program evaluation activities, current statistical methods for analysis of

these data are hardly used by program evaluators. With the availability of computer programs such as Everyman's Contingency Table Analyzer (ECTA) to perform the necessary calculations, log-linear and logit models can be applied to the analysis of cross-classifications to provide insights into complex evaluation data.

The purpose of this article is to show how the logit model can be used to explain variations in teacher attitudes. Specifically, the objective is to use data presented in Table 1 to analyze and predict classroom teachers' attitude toward proficiency tests based on knowledge of ethnicity, teaching grade level, and sex of the teachers. We have provided hypothetical data in Table 1 to illustrate the use of logit models for data analysis.

Table 1 is a four-way cross-classification of teachers by four dichotomous variables--ethnicity (1), sex (2), teaching grade level (3), and attitude toward proficiency tests (4). The rows are associated with three explanatory variables: ethnicity, sex, and teaching grade level. The columns refer to the dependent variable--positive or negative attitude toward proficiency tests. Both the conditional proportions (probabilities) and odds of expressing positive (and negative) attitude toward proficiency tests are also presented in the table. The data in the table show for example, that 51% of the 135 minority, male elementary school teachers have positive attitudes toward proficiency tests, while only 35% of white female secondary school teachers have positive attitude toward the tests.

A similar way of examining the same data is in terms of odds of expressing positive attitude toward proficiency tests. For minority male elementary school teachers, the odds are 69 to 66 (or 1.05 to 1) of expressing positive attitude toward the tests, while the corresponding odds for white female secondary school teachers are 50 to 92 (or 0.54 to 1).

The data in the table also show that the best predictor of attitude toward proficiency tests is ethnicity. Notice that, in general, a higher proportion of minority teachers expressed positive attitude toward the tests than white teachers, regardless of sex and teaching grade level.

The objective of our logit analysis is to build a simple and plausible model for the data presented in Table 1 and to estimate the effect of each explanatory variable on attitude toward proficiency tests. Our strategy would be to calculate the partial regression coefficients from the data, and to use them to answer the following questions:

1. What is the "main effect" of (a) ethnicity, (b) sex, and (c) teaching grade level on attitude toward proficiency tests?
2. Are there interaction effects? For example, does the effect of ethnicity on attitude toward proficiency tests depend on sex and/or teaching grade level?

TABLE 1

Cross-Classification of Teachers with Respect to
Ethnicity, Sex, Teaching Grade Level
and Attitude Toward Proficiency Tests:
Hypothetical Data

Ethnicity	Sex	Teaching Grade Level	Number of Teachers Expressing: Positive Attitude Frequency	Positive Attitude Probability	Positive Attitude Odds	Negative Attitude Frequency	Negative Attitude Probability	Negative Attitude Odds	Total
Minority	Male	Elementary	69	.51	1.05	66	.49	0.96	135
Minority	Male	Secondary	117	.64	1.80	65	.36	0.56	182
Minority	Female	Elementary	68	.38	0.61	112	.62	1.65	180
Minority	Female	Secondary	90	.52	1.08	83	.48	0.92	173
White	Male	Elementary	47	.34	0.51	93	.66	1.98	140
White	Male	Secondary	73	.48	0.91	80	.52	1.10	153
White	Female	Elementary	49	.25	0.33	149	.75	3.04	198
White	Female	Secondary	50	.35	0.54	92	.65	1.84	142
			563	.43	0.75	740	.57	1.33	1303

Conversion Formulas:
Odds = Prob/(1-prob)
Prob = Odds/(1+odds)

Independence Model: The independence logit model was proposed for the analysis of the data. The model tests the hypothesis that all the differences in the cell frequencies or the odds in Table 1 can be accounted for by "partial correlations constant in magnitude from condition to condition." In other words, the effects of the explanatory variables--ethnicity, sex, and level of teaching assignment--are mutually independent: There are no interactions among the variables and the probability of an observation falling in a particular cell of the table is the product of the marginal probabilities. As pointed out above, unlike other statistical methods, e.g., ANOVA and regression, the criterion to be analyzed in the logit model is the expected odds as a function of the explanatory variables.

The additive model is:

$$\text{Log } \Omega^{A}_{ijk} = \beta + \beta_i^{E} + \beta_j^{S} + \beta_k^{G} \qquad (1)$$

where Log Ω^{A} is the natural logarithm of omega, the expected value of the conditional odds of having a positive attitude toward proficiency tests.

β_i^{E}, β_j^{S} and β_k^{G} are the "main effects," that is the effects of the explanatory variables ethnicity, sex, and teaching grade level respectively.

$$E = \begin{cases} +1 & \text{if ethnicity is minority} \\ -1 & \text{White} \end{cases}$$

$$S = \begin{cases} +1 & \text{if sex is male} \\ -1 & \text{Female} \end{cases}$$

$$G = \begin{cases} +1 & \text{if teaching grade level is elementary} \\ -1 & \text{Secondary} \end{cases}$$

The model states that the expected odds, Log Ω^{A}_{ijk} pertaining to attitude toward proficiency tests (A) are dependent upon ethnicity (E), sex (S), and level of teaching assignment (G).

The saturated logit model for the same data is:

$$\text{Log } \Omega^{A}_{ijk} = \beta + \beta^{E}_{i} + \beta^{S}_{j} + \beta^{G}_{k} + \beta^{ES}_{ij} + \beta^{EG}_{ik} + \beta^{SG}_{jk} + \beta^{ESG}_{ijk} \qquad (2)$$

This model is called saturated because it contains as many parameters (main effects and interaction effects) as there are cells in the table. The saturated model fits the observed data perfectly because the expected frequencies estimated under it equal the observed frequencies. The saturated model presented above is unnecessarily complex as a description of the data, because it incorporates all the significant and non-significant relationships observed in the data presented in Table 1. Our objective, however, is to find a minimum set of meaningful patterns that account for the variations in the attitudes toward proficiency tests. The saturated model is presented, above, mainly to show how it was used to aid the search for the simplest and plausible model for the data.

Results

The ECTA Program was used to perform the analysis of the data. The results of the analysis have been presented in Table 2. The strategy adopted for selecting the simplest model to fit the data was to examine the standardized values of the effect parameters of the saturated model presented in Table 2. Column 2 of the table lists the λ (Lambda) effects of the parameters of the model. The λ effects are quantitative estimates of the direction and magnitude of the relationships between the variables in Column 1.

The standardized values of the Lambdas in Column 3 (Table 2) were obtained by dividing each estimated lambda effect by its estimated standard error. The standardized values can be used to test the statistical significance of a particular variable. A standardized value of 1.96, for example, indicates that the corresponding λ is statistically significant at .05 level.

As shown from Table 2, the interaction terms GSE, GSA, GEA, SEA, and GSEA are not statistically significant (each standardized values is less than 1.96). However, the following relations, GA, SA, and EA are statistically significant. In other words, teaching grade level, sex, and ethnicity contribute in a statistically significant way to the variation in the odds pertaining to teachers' attitudes toward proficiency tests. An examination of the estimates from the saturated model presented in Table 2 led to the hypothesized independence model (1). In the independence model, the lambda effects of the non-significant interactions GSE, GSA, GEA, SEA, and GSEA were set to zero.

TABLE 2

Estimates of the Main Effects and Interaction Effects of Ethnicity, Sex, and Teaching Grade Level on the Odds of a Teacher Expressing Positive Attitude Toward Proficiency Tests Under the Saturated Model

Variable	Lambda Effect (λ)	Standardized* Value of (λ)
A	-.143	-4.903
GS	-.080	-2.759
GE	-.042	-1.430
GA	-.139	-4.763
SE	-.014	- .477
SA	.125	4.293
EA	.169	5.823
GSE	.014	.472
GSA	-.003	- .112
GEA	-.002	- .065
SEA	.006	.220
GSEA	.008	.269

*Standard Error for λ effect is .029
The standardized value of λ = λ effect/standard error

The estimates of the effects of ethnicity, sex, and teaching grade level on teacher attitude toward proficiency tests for the independence model have been presented in Table 3. Each lambda effect in Column 2 of the table has been doubled to obtain the beta effects in Column 3. Also presented in Column 5 are the tau effects. Beta values are obtained from tau effects by using the relationship $\beta_i^E = 2 \text{ Log } T_i^E$.

It is noticed from the data in Table 3 that ethnicity, sex, and teaching grade level make statistically significant contributions to the variation of teacher attitudes toward proficiency tests. The betas show that ethnicity is the most important predictor of teacher attitude toward proficiency tests. Teaching grade level is next in importance followed by sex. Like the additive coefficients of ordinary regression, a positive beta effect indicates that the particular explanatory variable raises the odds on the dependent variable measure, while a negative beta indicates that the odds are decreased. Thus, being an elementary school teacher reduces positive attitude toward proficiency tests (-.278) while being a minority or a male increases it (.338 and .250 respectively).

TABLE 3

Estimates of the Main Effects of
Ethnicity, Sex, and Teaching Grade Level on
the Odds of a Teacher Expressing Positive Attitude Toward
Proficiency Tests Under the Independence Model

Variable	Lambda Effect (λ)	Beta Effect** (β)	Standardized* Value	Tau Effect (T)
A	-.142	-.284	-4.899	.867
G3	-.079	-.159	-2.701	.924
GE	-.041	-.082	-1.394	.960
GA	-.139	-.278	-4.772	.870
SE	-.016	-.032	- .540	.984
SA	.125	.250	4.308	1.133
EA	.169	.338	5.803	1.184
GSE	.011	.022	.382	1.011

*Standard error for λ effects is .029
**β effect = 2λ effect -- Both λ and β have the same standardized value

The beta effects in Table 3 can be used to reproduce exactly the expected odds from the independence model. For example, the expected odds for minority male elementary school teacher expressing a positive attitude toward the tests is 68.36 to 66.64 or 1.026 (see Table 4).

The equation for this logit is: $\text{Log } \Omega_1^A = \beta + \beta_1^E + \beta_1^S + \beta_1^G$ (3)

To show that the beta effects predict the expected odds of 1.026 for minority, male elementary school teacher, the appropriate betas (from Table 3) are plugged into the equation (3) to obtain:

$$\text{Log } \Omega_1^A = -.284 + .338 + .250 + (-.278) = .026$$

The natural anti-logarithm of .026 = 1.026. Thus the betas can be used to make specific predictions of the expected odds. The appropriate beta values in Table 3 have been used to make specific predictions of the expected odds of expressing positive attitude toward proficiency tests. The predictions have been presented in Column 6 of Table 4.

TABLE 4

Expected Cell Frequencies for the Independence Model

			Attitude		
Ethnicity	Sex	Teaching Grade Level	Positive	Negative	Specific Predictions Using Betas
Minority	Male	Elementary	68.36	66.64	1.03
Minority	Male	Secondary	116.70	65.30	1.79
Minority	Female	Elementary	69.00	111.00	0.62
Minority	Female	Secondary	89.94	83.06	1.08
White	Male	Elementary	48.04	91.96	0.52
White	Male	Secondary	72.90	80.10	0.91
White	Female	Elementary	47.60	150.40	0.32
White	Female	Secondary	50.57	91.53	0.55
			563.00	740.00	

Additional information that can be obtained from Table 3 concerns partial relationships, that is, the relationship between say ethnicity and attitude when the effects of sex and teaching grade level have been controlled or taken into account. Partial odds can be obtained from the data in Table 3 by simply raising the tau effects of the appropriate variables to the fourth power. Table 5 presents the information on the partial relationships between each explanatory variable and attitude toward proficiency tests. As shown in the table, minority teachers are about two times (exactly 1.965) more likely to have positive attitudes toward proficiency tests than their white counterparts teaching at the same level and of the same sex. Male teachers are also about 1.6 times more likely to have positive attitudes toward the tests than female teachers of the same ethnicity and teaching at the same level in the school system.

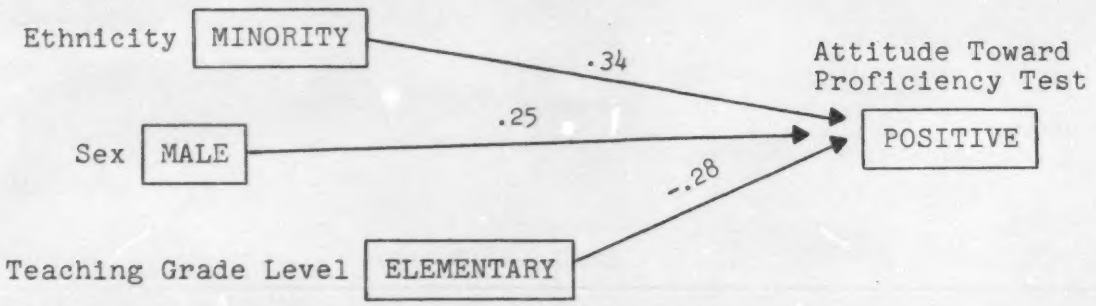
TABLE 5

Partial Relationships or Partial Expected Odd-Ratios
Pertaining to Attitude Toward Proficiency Tests

Variables	A Attitude (Mean)	E Ethnicity	S Sex	G Teaching Grade Level
Tau	.867	1.184	1.133	.870
Gamma (Tau-Squared)	.752	1.402	1.284	.757
Gamma-Squared		1.965	1.648	.573

The betas have also been presented graphically in a path diagram (Figure 1) to show how the explanatory variables affect attitude toward proficiency tests.

FIGURE 1



The relative effect of each independent variable on attitude can be measured by the difference between the betas. For example, ethnicity on the average improves attitude toward proficiency test more than sex.

Testing for Fit

Table 6 summarizes the results from ECTA for the saturated and the independence logit model. It is seen from the table that the independence model fits the data well as indicated by the small chi-square value of .13 with 4 degrees of freedom. Thus the proposed "main effects only" or the independence model provides a parsimonious explanation of the original data.

TABLE 6

Results from Saturated and Independence Models for the Prediction of Teacher Attitude Toward Proficiency Tests

Model	Explanatory Variables Included in the Model	Likelihood Ratio Chi-Square	Degrees of Freedom	P
Saturated	All (G, S, E, A plus all Interactions)	0	0	-
Independence	G, S, E, A (with no Interactions)	.13	4	>.5

Summary

We have discussed very briefly the use of the logit model for the prediction of a dichotomous variable based on three dichotomous explanatory variables. Using hypothetical data on teachers' attitude toward proficiency tests, we have shown how the odds of expressing positive attitude toward proficiency tests are effected by ethnicity, sex, and teaching grade level. We have provided indices of the magnitude of contribution made by ethnicity, sex, and teaching grade level to teachers' attitude toward proficiency tests.

REFERENCES

Davis, J. A. (1974) Hierachical models for significance tests in multivariate contingency tables: an exegesis of Goodman's recent papers. In Herbert Costner (Ed.), Sociological Methodology 1973-74, Jossey-Bass, San Francisco, pp. 189-231.

Goodman, L. A. (1979a) "A brief guide to the causal analysis of data from surveys." American Journal of Sociology 84: 1078-1095.

------, (1979b) "Multiplicative models for square contingency tables with ordered categories." Biometrika 66: 413-418.

------, (1979c) "Simple models for the analysis of association of cross-classifications having ordered categories." Journal of the American Statistical Association 74: 537-552.

------, (1979d) "Multiplicative models for the analysis of occupational mobility tables and other kinds of cross-classification tables." American Journal of Sociology 84: 804-819.

------, (1973a) "Causal analysis of data from panel studies and other kinds of surveys." American Journal of Sociology 78: 1135-1191.

------, (1973b) "The analysis of multidimensional contingency tables when some variables are posterior to others: a modified path analysis approach." Biometrika 60: 178-192.

------, (1972a) "A modified multiple regression approach to the analysis of dichotomous variables." American Sociological Review 37: 28-46.

------, (1972b) "A general model for the analysis of surveys." American Journal of Sociology 77: 1035-1086.

------, (1970) "The multivariate analysis of qualitative data: interactions among multiple classifications." Journal of the American Statistical Association 65: 226-256.

------, (1969) "On the statistical analysis of mobility tables." American Journal of Sociology 70: 564-585.

Knoke, D., and P. J. Burke (1980) Log-Linear Models, Beverly Hills, CA: Sage.

Piazza, T. (1975) A Working Guide to ECTA, Survey Research Center, U. C. Berkeley

Reynolds, H. T. (1977) Analysis of Nominal Data. Beverly Hills, CA: Sage.

------, (1977) The Analysis of Cross-classifications. New York: The Free Press.

Upton, Graham J. G. (1978) The Analysis of Cross-tabulated Data. New York: Wiley.